CENTER FOR INDEPENDENT EXPERTS (CIE) EVALUATION OF RIVER TEMPERATURE SUPPORT TOOLS

CALIFORNIA'S CENTRAL VALLEY

By

André St-Hilaire

Institut national de la recherche scientifique (INRS)
Centre Eau, Terre et Environnement
490 de la Couronne, Québec (Québec), CANADA G1K 9A9

December 2017

Reference

St-Hilaire, A. 2017. Evaluation of River Temperature Support Tools for California's Central Valley, iii+15 pages

TABLE OF CONTENT

1 Executive Summary

This report presents the author's independent peer review of the River Temperature decision support tools used to forecast water temperature for water and fisheries management in California's Central Valley. The assessment focused on the information gathered in the Technical Memorandum entitled:" Calibration and Validation of Linked Water Temperature Models for the Shasta Reservoir and the Sacramento River". Other documents, including related peer-reviewed papers were also consulted for this review. To the author's knowledge, the documents provided are the best scientific information available on this topic.

It was found that the selected models are appropriate tools for the stated objectives. Two models are deterministic and one is statistical. The latter model has known limitations, and should be presented as a simple AR model. CE-QUAL-W2 is the model used to simulate and forecast water temperatures in the Shasta Reservoir. It accounts for longitudinal and vertical thermal variability. Some biases in the model were highlighted and missing information (on the method used for selecting the discretization level, goodness of fit metrics, and the potential impact of linear interpolation) should be added.

A simplified version of the RAFT model was used for river temperature forecasting. Again, model biases were highlighted and the paucity of information on the impact of a selection of some parameters (only 3) for calibration is also noted.

The final linked model cascade is adequate and useful for the stated objectives, albeit plagued with important biases. Recommendations include 1) investigating the merits of other calibration approaches or 2) proceeding with post hoc correction of the biases.

2 BACKGROUND

Thermal models have become important management tools for management of water releases downstream of dams. In most cases, water release scenarios are designed to maintain key fish habitat features during their different life stages in rivers. In California's Central Valley, there are important reservoirs, including the Shasta Reservoir that regulate in part the Sacramento River system. The presence of the dam has potential adverse effects on winter-run Chinook salmon spawning, egg incubation and fry emergence. The most likely impact is associated with warm water temperatures that are sub-optimal for these critical life stages. In order to better manage water releases and mitigate the impact of warm waters downstream of the reservoirs, a model cascade has been implemented to simulate/forecast water temperatures in the reservoirs and in the river reaches impacted by water releases from the Shasta reservoir. The present report provides a review of the modelling tools being implemented. This review focuses on the modelling tools, highlighting their potential strengths and potential improvements, as this is the main area of expertise of the author.

3 STRENGTH AND WEAKNESSES OF INDIVIDUAL MODELS AND THEIR LINKS

3.1 Reservoir Model: CE-QUAL-W2

The model used to simulate temperatures in the reservoir is CE-QUAL-W2, which is a two-dimensional (longitudinal and vertical), Hydrodynamic and Water Quality Model (Cole, 2006). This model, created by the U.S. Army Corps of Engineering, has evolved in numerous iterations for over 40 years. The version of the model that is used is most likely the last (more recent) one, but this should be specified in the reports and web site.

The model discretizes the reservoir in layers along the main (longitudinal) axis and uses conservation of mass and momentum longitudinally and vertically. An overview of the geometry (number and size of units) used to discretize the Shasta reservoir is given in Figure 5. It is unclear if different configurations were tested. However, a vertical resolution of 1.5 m and 63 longitudinal segments are likely sufficient to capture the longitudinal and vertical thermal variability in the reservoir.

Initialization of the model requires that initial temperature on the entire domain be specified. This is done using data from one vertical profile for each initialization. It is not clear from the reviewed literature if a spin up time is subsequently required (most likely) and if this is the case, how long it has to be. Moreover, it is stated that temperatures were linearly interpolated to match the vertical grid resolution. Vertical profiles of temperature in reservoirs are most often non-linear. This approach may induce a certain amount of error in the initialization process. Some goodness-of-fit metric could be calculated or minimally, a figure showing some of the profiles and the interpolated temperatures would be useful.

Meteorological inputs used for the W2 model are described in Section 4.1.2. It is mentioned that data were downscaled to hourly time step using linear interpolation. The original time step, prior to downscaling, should be mentioned. Daily temperatures are non-linear and linear interpolation may introduce a bias if the time increment is long (e.g. more than 3 hours). This non-linearity was accounted for in the data gap-filling step, where singular spectrum analysis was used. Non linearity is even more important for precipitation.

Leakage was deemed important and was accounted for in the model structure. It is argued that including leakage leads to major model improvement, according to Figure 9. However, this figure compares simulated temperatures for the year 2012 only. Leaked flows or volumes are not reported in the main report (the authors cite a consultant report), but looking at Figure 12, these leakages appear to be important in the spring and fall, but not during the warmer period. Further explanations would be required.

The model calibration was performed using a split sample approach (odd years for validation and even years for calibration). The fact that years of calibration were not continuous implies that there was little if any provision for a spin up period. Was this a problem? Also, were the calibration years representative of the full climate and hydrological variability encountered in the past? Is this a limitation? A Bayesian approach was used to calibrate the model, which yields a full posterior distribution of each calibration parameter. This is very useful information, as it shows which parameter value is the most probably adequate. It is not clear, however, how this information is subsequently put to use. Was the final parameter value only the most probable one (50^{th} percentile)? Were different values of the distribution used? Furthermore, validation was based on a so-called "model skill", which appears to be a simple bias calculation. If it is the case, this should be specified. What about variance? A RMSE is a joint measure of bias and variance and as such, may be more adequate to assess model performance.

Figure 14 shows observed vs. predicted temperatures. RMSE are quite low, as well as bias. During the warmer months of July and August, the model seems to be slightly more biased, which is acknowledged in the report. However, this bias appears to be caused mostly by an underestimation of the warmer temperatures. Calculation of a bias for temperatures above 20 °C would be of interest, given the importance of high temperatures for fish.

3.2 Reservoir model: ARIMA model for Keswick Reservoir

It is explained in the report that because of the lack of geometry data for the Keswick Reservoir, an ARIMA model was used. In fact, given that the ARIMA (1, 0, 0) model was used, it is a simple AR model. Reported biases are relatively low, so are RMSE. However, there is no mention of the potential challenge of AR models in forecasting, especially in long term forecasting. It would be important to highlight the importance of data assimilation for such models, because of their auto-regressive component.

3.3 River model: RAFT

RAFT is a one-dimensional hydraulic model. Its spatial resolution is 2 km and its temporal resolution is 15 min. Some information on the process that culminated in the selection of these space and time steps would be useful. For instance, is there loss of information on potential lateral variability in some of the river reaches or is the water well mixed throughout the river? Is a 15 min time step truly required?

The RAFT model used is a simplified version of the original model created by NOAA. Simplifications include linear upstream-downstream interpolation of flows, rather than using the original Muskingum flow routing routine. The advection-dispersion model was also simplified. It is stated that these computational simplifications lead to small reductions in accuracy compared to errors associated with meteorological input data and the reader is referred to reference [7]. It would be useful to further quantify this loss of accuracy and perhaps contrast it to the gain in computation time, which is the most probable reason for making these simplifications (although this is not specifically stated in this section).

Data assimilation using a Kalman filter is implemented when Raft is in forecast mode. Initial assumptions about the distribution of error are required when using a Kalman filter. There is no information on what those assumptions were for RAFT data assimilation. This would be useful for a more thorough understanding of the final error distribution.

Some sort of sensitivity analysis on RAFT parameters was performed. It is stated that three parameters were “selected” to calibrate the model. What about the other parameters? Were they fixed? How were the values selected?

Overall, the RAFT model performance (RMSE and Bias) is good. The systematic underestimation of peak temperatures was duly noted in the report. In forecast mode, this bias should be corrected.

3.4 Linked models

The linked models appear to perform well overall, with RMSE ≤ 1.23 °C. However, although the mean bias is reported to be -0.1 °C, it is recognized that during the summer months the bias can be much higher. In fact, although the equation used to calculate bias is not shown, it is likely not using absolute values and the low bias is probably caused by the fact that linked model consistently underestimates low values and over estimates high values. Thus positive and negative biases compensate each other in the mean. The fact that the linked model overestimates maximum temperatures by 1°C or more during the warm period is troublesome for water release management.

The conclusion of the technical memorandum clearly states the usefulness of the model and highlights some of its limitation. I would like to draw attention to the fact that model biases can be overcome. One clear alternate approach would be to calibrate the model against more than one goodness of fit metric. For instance, in hydrology, it is well known that the Nash coefficient biases the calibration towards a better fit on floods than low flows. A similar approach can be used to ensure that the model is less bias on the temperature metrics that matter for salmon. For instance, degree-days metrics could be used in the calibration process. Another approach would be to proceed with post hoc debiasing. Bias could be eliminated using a quantile-quantile approach such as the one used for meteorological data. Finally, given that a full distribution of parameters is provided by the Bayesian calibration approach, an ensemble forecast could be used and the selected parameter quantiles could assist accounting for known biases.

4 INCLUSION OF UNCERTAINTY

As stated in the previous section, uncertainty was accounted for in many steps of the model cascade. Meteorological data were bias corrected. In addition, different percentiles of the GEFS variables were used as inputs for forecasting, thereby encompassing most of the distribution of the input variables. There is no clear statement on the impact of initial conditions on model uncertainty and on the requirement (or not) of a spin-up period for the model. In W2, leakages were shown to be important and were accounted for in the model to reduce uncertainty. The question of model calibration is addressed in the previous section and recommendations were made that could improve the quantification on uncertainty.

The ARMIA (AR) model used in the Keswick Reservoir may also impart large errors in forecast mode, given its autoregressive nature. If the forecast horizon is long, there is a chance that the model will drift, as it relies on previous forecasts for its current temperature estimation. Data assimilation is therefore of the utmost importance.

A Kalman filter is used for data assimilation. Kalman filters account for the distribution of errors in the data assimilation process, an important improvement from deterministic data assimilation. Comments on the error distribution were made in the previous section. An alternative approach to Kalman filters that does not require any assumption on a priori distribution of error is the Particle Filter approach (Arulampadam, 2002). The relevance of this approach in the context of data assimilation in this modelling exercise could be investigated.

Further recommendations on an Ensemble forecast approach were made in the previous section. Again, such an approach may provide more insight on the propagation of errors and the error imparted by model calibration.

5 IMPLICATION OF THIS WORK AS DECISION SUPPORT TOOLS

The model cascade has a strong potential to assist in decisions related to cold water releases. However, as stated previously, this potential could be enhanced if further efforts are made to minimize the bias, which can be quite large. Some potential approaches were listed in Section 4. Other approaches to minimize bias can likely be entertained.

6 CONTENT OF CVTEMP WEBSITE

The CVTEMP website is well structured. The section entitled “About CVtemp” could be expanded. The interactive maps are always key for this type of work. In the modelling section, it is interesting to see that the VIC model is being implemented. This will be a welcome addition to the model Cascade, but will come with its load of uncertainty. The interactive graph with flows and water temperatures is very informative, with histroric and current states being included.

The reservoir model section is also well structured, although more information about model design, discretization and mathematical structure could be provided. The same comment is true for the river model structure. Information about the ARIMA (or AR) model is non-existent.

The biological information is included in a very interesting graphic presentation of the so-called survival landscape, which is of the utmost interest to all stakeholders. Stations can be identified using horizontal lines on the graph. This is a nice feature. How is the probability survival contours button supposed to work? Showing Redd locations is also a very strong feature of this graphical display.

Meteorological inputs are also shown graphically. They should perhaps be moved further up in the suite of options. Time series can be downloaded for further analysis and key references are presented. This is therefore a very complete website.

7 CONCLUSION

A review of the River Temperature Decision Support Tools for California’s Central Valley was conducted and key points are summarized here. The final linked model cascade is adequate and useful for the stated objectives, albeit plagued with important biases in some instances. Recommendations include 1) investigating the merits of other calibration approaches or 2) proceeding with post hoc correction of the biases; 3) investigate the merit of alternative data assimilation approaches and ensemble forecasts.

8 REFERENCES

Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp, T. (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on signal processing 50(2), 174-188.

Cole, T.M., and S. A. Wells (2003). "CE-QUAL-W2: A two-dimensional, laterally averaged, Hydrodynamic and Water Quality Model, Version 3.1," Instruction Report EL-03-1, US Army Engineering and Research Development Center, Vicksburg, MS.

9 STATEMENT OF WORK

Statement of Work
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

River Temperature Decision Support Tools

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The SWFSC Fisheries Ecology Division (FED) requests an independent review of the suite of temperature modeling tools they have developed for water and fisheries management in California's Central Valley. When Shasta Dam was built in the 1940s it blocked Sacramento River Winter-run Chinook (SRWRC) salmon from accessing the cold waters of their native spawning habitat. The quality (water flow and temperature) of their current habitat below the dam is now entirely controlled by releases from the dam, and because SRWRC are listed under the Endangered Species Act, dam operations must take into account the impacts on their spawning and rearing habitat. As a result, temperature compliance points have been established:

> From Reclamation's 2008 OCAP Biological Assessment, Chapter 2, pg. 2- 38 (http://www.usbr.gov/mp/cvo/ocap_page.html):
>
> "In 1990 and 1991, SWRCB issued Water Rights Orders 90-05 and 91-01 modifying Reclamation's water rights for the Sacramento River. The orders stated that Reclamation shall operate Keswick and Shasta Dams and the Spring Creek Power Plant to meet a daily average water temperature of 56°F as far downstream in the Sacramento River as practicable during periods when higher temperature would be harmful to fisheries. The optimal control point is the Red Bluff Pumping Plant. Under the orders, the water

temperature compliance point may be modified when the objective cannot be met at Red Bluff Pumping Plant."

Page 590 of the 2009 OCAP Biological Opinion starts off with RPA Action Requirements: http://www.westcoast.fisheries.noaa.gov/central_valley/water_operations/ocap.html

To aid in the water and fisheries management decisions, SWFSC has developed linked temperature models for the Shasta Reservoir and the Sacramento River to model how operations will impact water temperatures within the SRWRC spawning habitat. The SWFSC then developed a thermal tolerance model for SRWRC eggs (the most temperature sensitive life stage) and linked it to the temperature model. The combined suite of models allows for water and fisheries managers to evaluate how proposed seasonal water operations impact SRWRC eggs in a spatiotemporally explicit manner.

Requirements

NMFS requires three reviewers to conduct an impartial and independent peer review in accordance with the Statement of Work SoW, OMB Guidelines, and the Terms of Reference (ToR) below. The reviewers shall have working knowledge and recent experience in temperature modeling, with specific emphasis on water temperature modeling in both lentic and lotic fresh water systems (i.e. river and reservoirs), thermal performance modeling of ectothermic organisms with an emphasis on early life stage development in relation to temperature exposure, and experience linking physical and biological models. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for reviewers

Each CIE reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: Review the following background materials and reports prior to the review. The contractor will provide these documents (via electronic mail or made available at an FTP site) to the CIE reviewers.

M. Daniels, E. Danner. 2017. Technical Memorandum: Calibration and Validation of Water Temperature Models for the Shasta/Sacramento System.

Martin, B. T., A. Pike, S. N. John, N. Hamda, J. Roberts, S. T. Lindley, and E. M. Danner. 2017. Phenomenological vs. biophysical models of thermal stress in aquatic eggs. Ecology Letters. DOI: 10.1111/ele.12705

Pike, A., E. Danner, D. Boughton, F. Melton, R. Nemani, B. Rajagopalan, and S. Lindley. 2013. Forecasting river temperatures in real time using a stochastic dynamics approach. Water Resources Research 49(9):5168-5182. DOI: 10.1002/wrcr.20389

Danner, E. M., F. S. Melton, A. Pike, H. Hashimoto, A. Michaelis, B. Rajagopalan, J. Caldwell, L. DeWitt, S. Lindley, and R. R. Nemani. 2012. River Temperature Forecasting: a Coupled-Modeling Framework for Management of River Habitat. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(6):1752-1760. DOI: 10.1109/JSTARS.2012.2229968

The Central Valley Temperature Mapping and Prediction (CVTEMP) website: http://oceanview.pfeg.noaa.gov/CVTEMP/

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Place of Performance
Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Period of Performance
The period of performance shall be from the time of award through November 2017. Each reviewer's duties shall not exceed 10 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Within two weeks of award	Contractor selects and confirms reviewers
Within four weeks of award	Contractor provides the pre-review documents to the reviewers
October 2017	Each reviewer conducts an independent peer review as a desk review
Within two weeks after review	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each ToR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
Since this is a desk review travel is neither required nor authorized for this contract.

RESTRICTED OR LIMITED USE OF DATA
The contractors may be required to sign and adhere to a non-disclosure agreement.

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

3. The reviewer report shall include the following appendices:
    a. Appendix 1: Bibliography of materials provided for review
    b. Appendix 2: A copy of the CIE Statement of Work

Annex 2: Terms of Reference for the Peer Review

River Temperature Decision Support Tools

1. Evaluation of the strength and weaknesses of the individual water temperature models as well as the process of linking the models, bringing attention to those weaknesses not adequately addressed in technical memorandum.

2. Evaluation of the methods used to incorporate uncertainty into predicting water temperature in Shasta Reservoir and Sacramento River down to Red Bluff, such as the use of variable meteorology and model parameters.

3. Evaluation of the water temperature model calibration and validation procedure outlined in the technical memorandum and its ability to properly parameterize each water temperature model.

4. Evaluation of the implication of this work as decision support tools, bringing attention to the any potential for mis-use or mis-interpretation of this information to aid in fisheries and water management in California's Central Valley

5. Evaluation of the content made available in the CVTEMP website, bringing attention to content that was unclear and that could be improved.

6. Provide a brief description on other aspects of the model not described above.